Increased Exposure to Rigid Routines can Lead to Increased Challenging Behavior Following Changes to Those Routines

Bull, L. E., Oliver, C., Callaghan, E., & Woodcock, K. A. (2015). Increased Exposure to Rigid Routines can Lead to Increased Challenging Behavior Following Changes to Those Routines. *Journal of Autism and Developmental Disorders*, *45*(6), 1569-1578. https://doi.org/10.1007/s10803-014-2308-2

Published in:
Journal of Autism and Developmental Disorders

Document Version:
Peer reviewed version

Leah E. Bull, Chris Oliver, Eleanor Callaghan and Kate A. Woodcock, Cerebra Centre for Neurodevelopmental Disorders, University of Birmingham, UK.

Author Kate A. Woodcock is now at the School of Psychology, Queen's University Belfast, UK.

Acknowledgements. This work was supported by a project grant from the Jérôme Lejeune Foundation to KAW and CO, and Cerebra who provide core funding to the Cerebra Centre for Neurodevelopmental Disorders (to director CO). Special thanks to the Prader-Will Syndrome Association, UK and Gretton Homes for their assistance in recruitment; Emma Cross, Laura Heath-Jones Campbell, Victoria Johnson, Jessica Penhallow, Amy Perry and Helena Todd, for assistance with data collection and processing; and Prof. Tony Holland for advice on project design and assistance with recruitment. Finally, we are extremely grateful for the support of the participants, their families and caregivers, without whom the work would not have been possible.

Correspondence should be addressed to Dr. Kate Woodcock, School of Psychology, Queen's University Belfast, University Road, Belfast, Northern Ireland, BT7 1NN. Email papers@katewoodcock.com. Telephone +44 (0) 28 9097 4886

This is the final preprint version of the manuscript published in the *Journal of Autism and Developmental Disorders,* in press in November 2014
The final publication is available at Springer via http://dx.doi.org[insert DOI]

Increased exposure to rigid routines can lead to increased challenging behavior following changes to those routines

Abstract

Several neurodevelopmental disorders are associated with challenging behavior following changes to routines. Here individuals with Prader-Willi syndrome, who show elevated levels of this behavior, are examined to better understand how experience with a routine can affect challenging behavior following its disruption. Play based challenges exposed 16 participants to routines, which were either adhered to or changed. Temper outburst behaviors, heart rate and movement were measured. As participants were exposed to routines for longer before a change, more temper outburst behavior and increased emotional arousal was elicited by changes. Further research is important to understand whether intervention approaches that limit exposure to changes may benefit from the structured integration of flexibility to ensure that the opportunity for routine establishment is limited.

Key words: resistance to change; restricted preferences; preference for routine; challenging behavior; temper tantrums; Prader-Willi syndrome;

Introduction

Individuals with neurodevelopmental disorders including autism spectrum disorder, and several genetically defined disorders such as Prader-Willi syndrome, commonly show a strong preference for routine and predictability (Kuenssberg, Murray, Booth & McKenzie, 2014; Moss, Oliver, Arron, Burbidge & Berg, 2009). Importantly, this preference for predictability can manifest as challenging behavior following changes to routines or expectations (Gomot & Wicker, 2012; Furniss & Biswas, 2012; Richards, Oliver & Allen, 2010; Sabaratnam, Murthy, Wijeratne, Buckingham, Payne, 2003; Woodcock, Oliver & Humphreys, 2009a). In Prader-Willi syndrome, this resistance to change is particularly prevalent; and associated with temper outbursts, which have been measured in experimental settings by tracking outburst component behaviors (Oliver, Woodcock Humphreys, 2009; Woodcock, Oliver & Humphreys, 2011). Here, Prader-Willi syndrome is used as a model for understanding the dynamics of the association between changes to routines / expectations and specific profiles of challenging behavior, in this case temper outbursts. This work will inform a broader strategy for the development of intervention approaches targeting difficulties with change experienced by people with neurodevelopmental disorders.

The comparability of the resistance to change in people with Prader-Willi syndrome (PWS), to such behavior in individuals with other neurodevelopmental disorders is supported by research into its cognitive correlates. At a cognitive level, the preference for routine and predictability in people with PWS has been linked to a specific cognitive deficit in task set re-configuration; a component process of task switching / shifting (Woodcock, Oliver & Humphreys, 2009b; Woodcock, Oliver, Humphreys & Hansen., 2010); which is an important aspect of executive function (Miyake, Friedman, Emerson, Witzki, Howerter & Wager, 2000).

This relationship also appears to be present in boys with Fragile X syndrome, which has a distinct genetic aetiology (Woodcock et al., 2009b). Whilst there has been some debate on the issue (White, 2013), converging evidence suggests that individuals with autism spectrum disorders also show deficits in measures of shifting (Russo, Flanagan, Iarocci, Berringer, Zelazo & Burack, 2007). Importantly, performance on shifting tasks has been associated specifically with the repetitive /restricted preferences domain of autism spectrum behavior; a domain which comprises the preference for predictability (D'Cruz et al., 2013; Lopez et al., 2005). These data suggest that the preference for predictability observed across several neurodevelopmental disorders – even those with distinctly different causes and phenotypes – may be associated with the same cognitive features.

Prader-Willi Syndrome is a neurodevelopmental disorder caused by the absence of paternally derived genetic material in the q11.2-13 region of chromosome 15. There is a well characterized physical phenotype (Holm et al., 1993), alongside mild to moderate intellectual disability (Whittington et al., 2004). Temper outbursts are shown by upwards of 80% of people with disorder (Dimitropoulos et al. 2001; Walz & Benson, 2000); and a common trigger for these outbursts is change to routine or expectations (Woodcock et al., 2009a; Tunnicliffe, Woodcock, Bull, Oliver & Penhallow., 2014). Some aspects of the phenotypic behaviors evidenced by individuals with Prader-Willi syndrome have been reported to vary across different genotypes that can cause the syndrome (e.g. Butler, Whittington, Holland, Boer, Clarke & Webb, 2002). However, both of the primary genetic sub-types appear to show similar rates of temper outbursts linked to changes to routines or expectations (Woodcock, 2009).

Existing approaches that seek to address resistance to change in individuals with neurodevelopmental disorders frequently aim to increase advance planning and predictability

(Mesibov, Browder & Kirkland, 2002). The rationale behind such approaches comes from behavioral theory and involves – after having identified changes to routines / expectations as an antecedent for challenging behavior – manipulating the environment in such a way that the frequency of occurrences of antecedents for the behavior is reduced. However, these approaches often result in individuals being exposed to increased repetition of the same sequences of events i.e. routines.

In relating the increased repetition of sequences of events to cognitive theory, such repetition corresponds to infrequent, compared to frequent, required task switches. There is evidence to suggest that the nature of the cognitive demand imposed by switching is different depending on whether such switches occur frequently or infrequently. Thus, while infrequent switches place higher demands on task-set reconfiguration, more frequent switches place higher demands on task-set updating (Monsell & Mizon, 2006; Nessler, Freidman & Johnson., 2012). Further discussion of the intricacies of these specific components of switching is not pertinent here. Nevertheless, these data suggest that, at least in relation to the specific switching deficit in individuals with PWS and Fragile X syndrome (Woodcock et al., 2009b), less frequent switches may place greater demands on this deficient process. This is relevant because there are data to suggest that in some individuals, disruption of a routine that the individual has experienced repeated previous exposure to can trigger challenging behavior, where disruption to a routine to which the person has been recently introduced occasions no behavioral difficulty (Woodcock, Oliver & Humphreys, 2011).

If it is the case that increased exposure to routines results in increased difficulty following changes to these routines, then this would have important implications for the development of intervention strategies. It would imply that antecedent manipulation approaches, which aim to reduce the changes to expectations in people's environments, should also be sensitive to minimizing opportunities for routines to become established. The question also has important implications for potential early intervention approaches. Anecdotally, it has been reported that families of children with PWS who show little resistance to change, also appear to be those who report few opportunities for routines to become established during children's development (Woodcock et al., 2009a; 2011). Mice models have demonstrated that development in a varied environment, in which there is decreased exposure to the same stimuli and events, results in increased cognitive flexibility and reduced behavioral routines (Tanimura, Yang & Lewis, 2008). These data suggest that increased exposure to the same sequences of events from an early age could have important potentially negative implications for later cognitive and behavioral functioning.

The primary aim of the present study was to investigate the effect of increasing length of exposure to a routine on challenging behavior following changes to that routine. Importantly, because temper outburst behavior in people with PWS was used as a model for this investigation, it was also possible to investigate the impact of such repeated routines on the physiological correlates of this behavior. Temper outbursts are often defined in relation to associated increases in emotional arousal (Potegal, 2003). Consistent with this definition, temper outbursts in individuals with PWS comprise consistent behavioral indicators of increased emotional arousal (Tunicliffe et al., 2014; Oliver et al., 2009). Emotional arousal is associated with increased activation of the autonomic nervous system, which can be indexed by increases in heart rate (Ekman, Levenson & Friesen, 1983; Rainville, Bechara, Naqvi & Damasio, 2006; Fernandez, Pascual, Soler, Elices, Portella & Fernandez-Abascal, 2012). However, heart rate is heavily dependent on physical activity (Iellamo, 2001). Thus, here both heart rate and physical activity are measured in order to index changes in emotional arousal following changes to routines to which individuals have been exposed for different lengths of time.

It was hypothesized that in a sample of individuals with PWS, increased exposure to a routine will be associated with increased temper outburst component behaviors and increased emotional arousal following changes to that routine.

Method

Participants

Ethical approval was obtained from (*removed for blind review*). All adult participants and parents of children under 16 years provided informed consent. Children under 16 also provided their informed assent. Participants were recruited from the Prader-Willi Syndrome Association in the UK (PWSA-UK) and from a group of residential homes for adults with PWS. Parents and carers were interviewed via telephone to ascertain the antecedents, component behaviors and consequences associated with the temper outbursts they observe (see *supplementary materials* for the interview schedule). Only individuals who displayed temper outbursts as a result of change to routine or expectation (though not necessarily the only trigger) were recruited. Sixteen individuals took part (*Table 1*). Vineland Adaptive Behavior Scales (Sparrow et al. 2005) were conducted to assess participants' adaptive behavior level to facilitate comparison with previous and future research.

Table 1: Descriptive information on participants

N	16
Age range (years: months)	9:7 – 47:10
Mean age (SD); years: months	25:0 (13:9)
N per gender: males: females	12:4
N per genetic subtype: mUPD: deletion: unknown	6:2:8
VABS adaptive behavior: Range	25-95
VABS adaptive behavior: Mean (SD)	64.4 (17.92)
VABS Daily Living Skills age equivalent: Mean (SD) in years: months	7:7 (3: 2)

Measures

Change challenge games

Four table top games were designed. These games were all novel to participants (see *supplementary materials*); and allowed routines to be established during the course of play. As an illustrative example, one of the games involved choosing cards from a central pile based on rolling a die; selecting counters to discard based on the chosen card; and then discarding the card into a different pile, not to be used again during that round. Thus, one of the routines established was the separation of the already played cards from those still available, and a change to this routine was mixing of an already played card back into the pile of cards still in play.

Change challenges were presented in either Disrupt or Establish conditions. In the Establish condition, routines and/or expectations were followed as expected, thus providing participants exposure to the corresponding routines without change. In the Disrupt condition, up to five changes (mean: 4.8; SD: 0.65) were imposed on the corresponding routines/expectations (see *supplementary materials* for a full description of these).

Physiological recordings

Participants wore a heart rate monitor (Polar RS400; to measure heart rate) and an Actiwatch (AW4, CamNtech Ltd., Cambridge, United Kingdom), containing an accelerometer, which measures activity. The heart rate monitor was worn on a strap around the chest with a watch on the wrist, and the Actiwatch was worn on the participant's wrist. Heart rate was recorded in average beats per minute (bpm) every second. Activity was recorded as an activity count. The accelerometer in the Actiwatch produces an electric current when movement is detected and the change in voltage is measured as an activity count. Activity counts were recorded in epochs of ten seconds (extracted using Actiwatch Activity and Sleep Analysis 7, Version 7.28, CamNtech Ltd).

Procedure

The experimental design was within subjects: each participant engaged with all four activities, each presented during an Establish condition (which varied in duration across activities), and a Disrupt condition. Thus, the effect of increasing exposure to routine on response to change to a routine could be evaluated for each participant individually. The procedure is summarized in *Figure 1.* Participants were assessed during one day at their home by a single researcher. Participants were first taught how to play each of the four games during a familiarization period (14 - 36 minutes); with matched duration of exposure to each game for any single participant. The purpose of the familiarization period was to ensure that participants understood the rules of the game. During this period, the researcher did not mention winning in order to minimize participants' focus on trying to win.

Figure 1

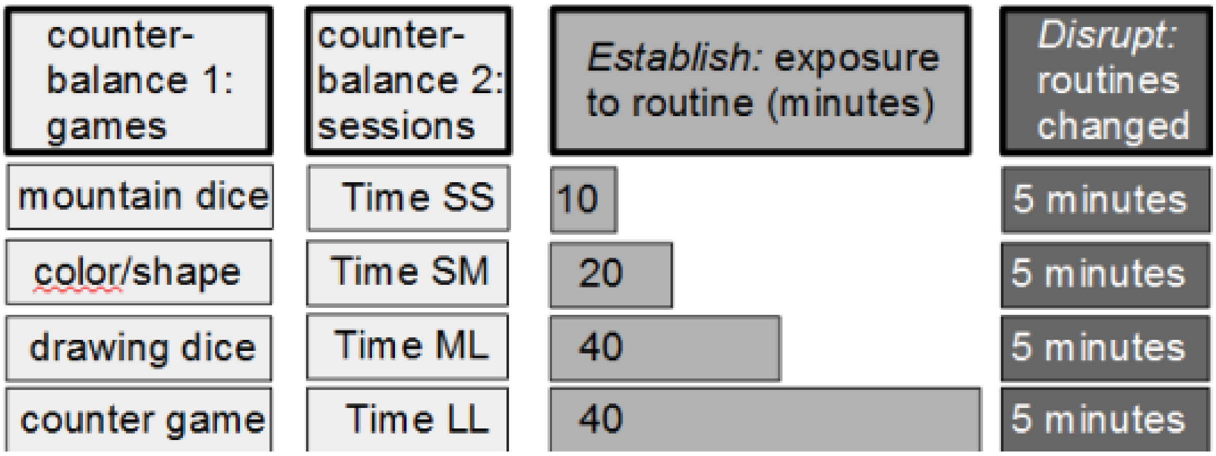

Figure 1 Experimental procedure. Time SS, SL, ML, LL are arbitrary labels for sessions comprising an Establish followed by a Distraction condition, which occurred at different times.

Participants then took part in Establish and Disrupt conditions in pairs, corresponding to every game. The Establish condition was always presented first; lasting either ten, twenty, forty or eighty minutes; and was immediately followed by the corresponding Disruption condition. Disruption conditions lasted at least five minutes. However, there was some variation across participants in the length of time required to explain or conduct the changes (mean duration: 7 minutes 23 seconds; SD: 2:51). Breaks were scheduled between each pair of conditions and no participants asked for breaks at any other time.

Importantly, two aspects of this procedure were fully counterbalanced across participants. Firstly, the game participants engaged in for each of the four possible durations of Establish condition; and secondly, the order with which games associated with each length of Establish condition, was presented to participants. This counterbalancing procedure minimized possible confounding effects of changes in motivation for play as the procedure progressed, and general habituation to changes being conducted by the researcher.

Behavior observation

Participants were filmed using a video camera whilst playing the games so that behaviors could be observed and analyzed. Behaviors of interest were temper outburst related behaviors that parents or carers had identified during the interviews (see *Participants*). Behaviors were coded in real time using ObsWin 3.2 (Martin, Oliver & Hall., 1998) based on operationally defined categories (e.g. Oliver et al., 2009) for which inter-rater reliability; based on two researchers coding 25% of each participant's data; demonstrated a Kappa value of at least 0.6 across 5 second time periods (*Table* 2).

Table 2: Definitions and reliability of observed temper outburst behaviors

Behavior	Operational Definition	Inter-rater reliability: Kappa
Questioning	The participant asks the researcher a question related to the game. These could be about the rules/materials/turns.	0.74
Ignoring Requests	The participant does not respond to a verbal request made by the researcher or the participant starts to verbalize about something unrelated to the request. This should be coded until a further verbal response from the researcher (either a further request or a verbalization about something unrelated to the request) or the participant stops ignoring and initiates a response.	0.88
Arguing	The participant makes verbalizations in the form of statements of disagreement, giving orders or making demands, taken from Oliver et al. (2009).	0.85
Crying	The participant shows tears or speech or non-speech vocalizations associated with crying, taken from Oliver et al. (2009).	0.96
Physical Aggression	The participant responds with a deliberate act towards researcher or object involving contact that could cause harm or damage. This should also include any missed attempts at physical aggression where no contact is made.	0.84
Verbal Aggression	The participant verbalizes threats or makes hurtful comments towards the researcher. This could also include any offensive language.	0.97
Gestural Aggression	The participant displays a behavior that can be viewed as threatening but involves no contact with the researcher or object, for example pointing.	0.93
Picking Nose	This additional behavior was coded for one participant only as this had been identified by their parents to be a temper outburst behavior. The participant engages in picking nose with fingers or tissue and includes blowing nose and includes eating any mucus from fingers or tissue.	0.69

Analyses

Analyses were based on mean percentages of time in which temper outburst behaviors were shown, mean heart rates, and mean activity counts; within the relevant conditions (i.e. Disrupt and Establish for the 10, 20, 40 and 80 minute routine exposure phases respectively).

Initial inspection of the observational data revealed distributions that significantly departed from normality (Kolmogorov-Smirnov statistic up to: .4; with $p < .001$). Thus, non-parametric analyses were employed. Firstly, as an assessment of experimental integrity, a Wilcoxon Signed Rank test was applied (using IBM SPSS Statistics 20 software) to assess the difference in temper outburst behavior in Disrupt relative to Establish conditions. Secondly, the effect of increased length of exposure to a routine on temper outburst behavior following change to that routine was assessed using a Page's Trend test. The Page's Trend test provides a non-parametric alternative to repeated measures Analyses of Variance (ANOVA). Importantly, the approach allows a hypothesis to be tested where the order of the treatments can be predicted (i.e. that more temper outburst behavior will be demonstrated in Disrupt conditions following longer Establish conditions), but the size of the difference between each of the ordered treatments cannot be predicted (i.e. there is no reason to predict that there would be a linear effect of increasing duration of Establish conditions on the temper outburst behavior demonstrated during corresponding Disrupt conditions). The Page's Trend test therefore provides greater statistical power for the present purpose relative to alternative approaches such as the Freidman's test (Page, 1963). The test was calculated manually using guidance from Meddis (1975) via the computation of Z scores that provide a measure of effect size in standard deviation units.

One child was not willing to complete more than part of the Establish condition for the first game and thus was not included in the analysis. Two adults did not wish to start/complete the game associated with an 80 minute Establish condition; thus these observational data were treated as missing. The missing data were dealt with in a conservative manner by taking the mean of temper outburst behaviors across all other corresponding conditions (i.e. Establish or Disrupt) for the relevant participant. Thus, the value substituted for the missing data could not strengthen the hypothesized effect if it were present (no effect on Type I errors), but dealing with the missing data in this way allowed the power of the test to be maximized (decreasing the likelihood of Type II errors).

Analyses of physiological data focused on a subset of 10 participants for whom full heart rate and Actiwatch data were available. Full data were not available for six participants, either because these individuals were not comfortable with wearing the recording equipment or due to technical failure of the recording devices. Only Disrupt conditions were assessed because the relatively long duration of Establish conditions meant that they were highly subject to effects of movement, which would confound differences in heart rate linked to physiological arousal. Because the first change was not imposed immediately upon initiation of Disrupt conditions, data were averaged for each condition over only the middle 80% of the time period of that condition. The distributions of the resulting mean activity and heart rate values did not significantly depart from normality (Kolmogorov-Smirnov <.211; $p > .200$). Thus, parametric analyses were conducted because it was necessary to assess both heart rate and activity data to inform on physiological arousal, and the assumed interplay between these two measurements meant that clear apriori directional hypotheses for both measures (as would be required for a Page's Trend test) could not be made. Thus, repeated measures ANOVAs with a single duration factor, comprising 10, 20, 40 and 80 minute levels, were applied to assess the effect of increasing length of exposure to a routine on heart rate and physical activity following change to that routine.

Results

Observational data

Supporting the experimental integrity of the present methods, the mean percentage of time during which temper outburst behaviors were presented across all games was significantly higher during Disrupt conditions relative to Establish conditions (Median Change: 3.31; median No Change: 1.43; Wilcoxon signed rank standardized value: 3.12, p = .002; Cliffs d=0.43).

In line with our hypothesis, there was a significant main effect of increasing Establish condition duration on the percentage of time during which temper outburst behavior was demonstrated in corresponding Disrupt conditions (L= 395, p = .038; Z=1.79; *Figure 2*).

Figure 2

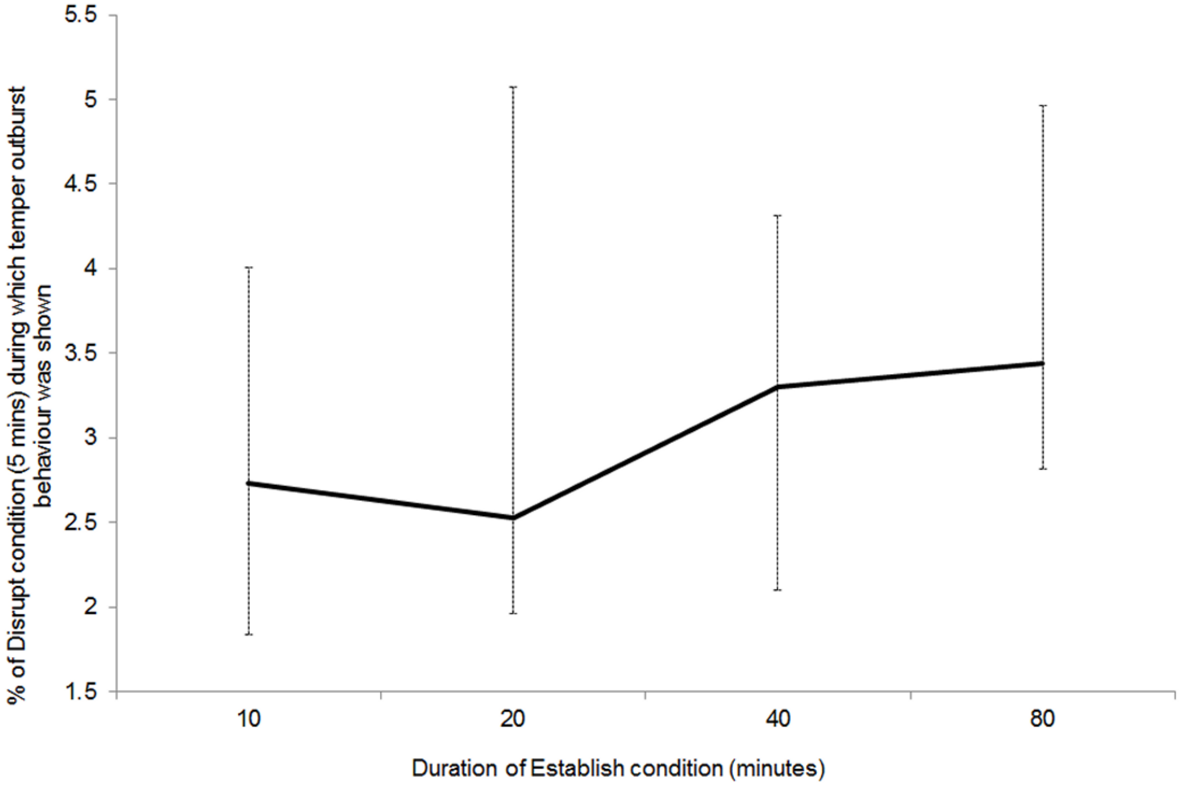

Figure 2 The median percentage of five minute Disrupt conditions during which temper outburst behaviors were demonstrated. Horizontal dashed lines represent the interquartile range.

However, inspection of the observational data revealed high levels of individual variability (*see supplementary materials for individual participant level data*). Further exploratory analyses revealed that an important factor contributing to the individual variability was the proportion of time participants spent distracted from the game (i.e. not looking at the researcher or the game; or talking about an unrelated topic) during Disrupt conditions. When participants who evidenced higher levels of distraction (20% or more of at least one Distract condition) were removed from the analysis (remaining n=9), the main effect of duration was stronger (*L*=244, p = .029, Z = 2.19); however this effect was not present in participants showing higher levels of distraction (p > .90). Overall, participants presently labelled as more distracted demonstrated more temper outburst behavior than those labelled as less distracted, but this effect only bordered significance (p= .066). Further, across participants, there was a significant association between increased duration of time distracted and increased total duration of temper outburst behavior (Spearman's r = 0.52, p = .050). The *supplementary materials* include additional details on these exploratory analyses.

Physiological data

Mean physiological measurements across relevant conditions are described in *Table 3.* The repeated measures ANOVA of heart rate data revealed a strong, significant main effect of duration ($F(3,27) = 3.13$, $p = .042$, $\eta^2_p = .26$). As illustrated in *Figure 3,* heart rate was higher in Disrupt conditions associated with longer Establish conditions, relative to that associated with the 10 minute Establish condition. However, after applying a Bonferroni correction to the (one-tailed) directional paired comparisons (adjusted threshold: $p < .017$), it was only the increase in heart rate during the Disrupt condition associated with the 20 relative to the 10 minute Establish condition, which attained significance ($t(9)=3.35$, $p = .008$, $\eta^2 = .55$). These results are also reflected in the linear increase in heart rate across Disrupt conditions associated with increasing length of Establish conditions, which was of medium size but did not attain significance ($F(1,9) = .77$, $p = .404$, $\eta^2_p = .078$). The quadratic main effect was strong and significant ($F(1,9) = 6.86$, $p = .028$, $\eta^2_p = .43$) but this was driven by the larger increase in heart rate in the Disrupt condition following the 20 minute relative to the 10 minute Establish condition. To assess how far these changes in heart rate could be explained by changes in physical activity, the activity data were assessed in the same way. Here, the repeated measures ANOVA revealed no significant main effect of duration ($F(3,27) = .89$, p = .46, η^2_p= .090); nor linear change in activity across Disrupt conditions (negligible linear effect; $F(1,9) = .006$, p = .940, η^2_p= .001). Thus, whilst in general increased duration of Establish conditions was associated with increased heart rate in corresponding Disrupt conditions, this relationship was not linear. Changes in physical activity did not appear to drive the relationship.

Table 3: Mean heart rate and activity counts for Disrupt conditions

Mean physiological measurement during Disrupt condition		Duration of associated Establish condition			
		10	20	40	80
HR (bpm)	Mean	74.28	81.22	78.83	76.41
	STD	13.51	16.73	16.38	13.67
Activity (count)	Mean	56.66	69.59	55.63	60.73
	STD	45.46	43.28	27.45	2.538

Figure 3

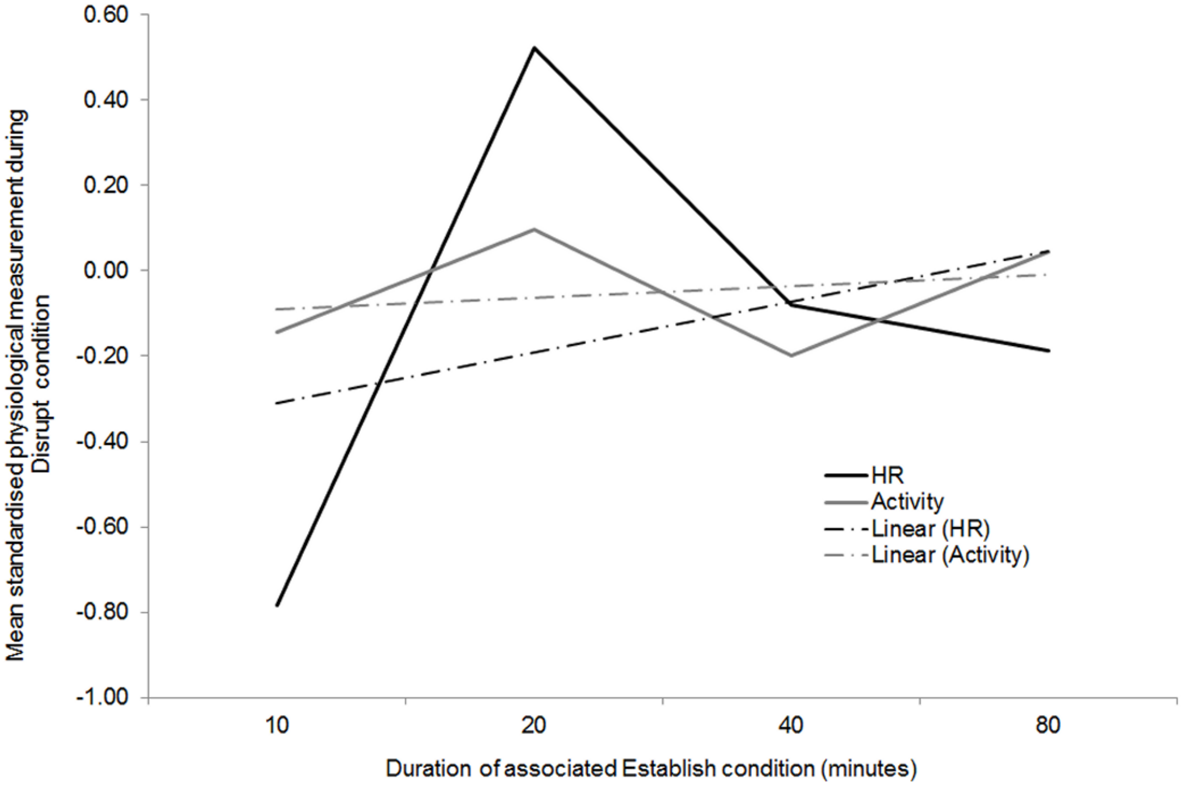

Figure 3 Heart rate and activity data in standardized units calculated based on the mean and standard deviation of recordings across all (Disrupt and Establish) conditions for each individual.

Discussion

Using a model of temper outburst behavior in individuals with Prader-Willi syndrome, the results provided support for the hypothesis that increasing exposure to a routine without change would be associated with increased behavioral difficulties following change to that routine. The physiological data provided additional support for this relationship because corresponding increases in heart rate were not driven by increases in physical activity; suggesting that these may have been underpinned by emotional arousal, which has been linked to temper outbursts. However, the effects of increased opportunity for the establishment of routines could not be described with a clear dose-exposure function, emphasizing the need for further research in this area. Our exploratory analyses highlighted additional factors that may impact on the relationship.

To the best of our knowledge, no previous work has directly examined the interaction between rigidity versus flexibility in individuals' environments during development and the behaviors that may be shown by these individuals following changes to routines later in life. However, an association has been demonstrated between rigidity in parent behavior during parent-child interactions and current and subsequent externalizing behavior in children (Hollenstein, Granic, Stoolmiller & Snyder, 2004) and this work has formed the basis of several studies examining the impact of maternal depression on children's externalizing behavior (e.g. Lunkenheimer, Albrecht & Kemp, 2013). In addition, obsessive-compulsive completeness traits – including a preoccupation with things being done in a particular, preferred way – are demonstrated significantly more frequently by parents of children with autism compared to parents of typically developing children; and increasing parental completeness is associated with increased resistance to change behavior in the children (Kloosterman, Summerfeldt, Parker & Holden, 2013). The causal direction of these relationships is unknown and parent and child behavior is likely to interact at several levels. However, taken together with the present findings, these data highlight an important need for further examination of the effects of increasing environmental rigidity during development on subsequent resistance to change and challenging behaviors following change. A cautionary note here is warranted because there are multiple reasons why increasing structure in children's environments is widely considered best practice in the context of several neurodevelopmental disorders, and the present results do not suggest any contraindication to this approach. In future, careful prospective designs, which work with the variation in standard advice proportioned to the families of individuals with different disorders or in different settings, as well as with individual differences in relevant family characteristics, are needed.

The present findings did not demonstrate a clear exposure-response relationship in the effects on challenging behavior of increasing exposure to routines. Whilst such a relationship was not specifically predicted, its characterization would have provided stronger support for a pivotal role of such exposure to routines. However, findings from several areas (e.g. the impact of environmental risks on children's externalizing behavior) also demonstrate robust associations between factors under investigation and behaviors, in the absence of exposure-response relationships (Donkin et al., 2013; Fraser, Kirkby, Daniels, Gillroy & Montgomery, 2001; Liu, Leung, McCauley, Ai, Pinto-Martin, 2013; Sinha, Manhar Husain, 2013). Some of the complexity of these relationships is likely to arise because of the multi-level factors that influence behavior. However, also particularly relevant to the present study, is the arbitrary nature of the exposure levels that may be contrasted (e.g. Fraser et al., 2001). In the present study, different durations of exposure to routines were selected on the basis of pragmatic concerns about procedural feasibility. It is possible that there was not enough variability between each pair of durations for a dose-exposure relationship to be identified. Future studies in this area, which contrast routines that have been established over several days, weeks or even months, would be informative in this respect.

Our exploratory analyses identified a behavioral category, labeled here as *distraction*, as an additional factor that can impact the relationship between exposure to a routine and the behavioral response to changes to that routine. Distraction, as defined here, comprised times when participants were not paying active attention to the activities. The effect of increasing exposure to routines remained present, and was stronger, in participants who demonstrated little distraction. However, the effect was not present in those participants who demonstrated relatively high levels of distraction.

One possible conceptualization of distraction is that it indexes times when participants were not “on-task”. On-task behavior has been closely linked to effective learning in educational settings and appears to be key to individuals benefiting from the specific features of carefully designed learning environments (Imeraj et al., 2013; Ponitz, Rim-Kaufmann, Grimm, Curby, 2009). Thus, it is possible that individuals who demonstrated high levels of distraction were simply less sensitive to the experimental manipulation. However, this explanation is not consistent with the finding that participants who spent more time distracted also showed more temper outburst behaviors. An alternative conceptualization of distraction is that it represented active attempts at emotion regulation by participants. Such self-distraction has been identified as a strategy, which whilst commonly shown by typical children, varies greatly across individuals in its efficacy for reducing negative emotions (Buss & Goldsmith, 1998; Ekas et al. 2011). Thus, one interesting possibility is that the present participants (to differing degrees) engaged in, but were not able to successfully manage their negative emotions using, a self-distraction strategy. These data on distraction are exploratory. However, importantly, they suggest that even in a group of participants with the same genetic disorder, recruited to show a specific pattern of challenging behavior in certain environmental circumstances, differences in how environmental challenges are managed across individuals may still result in different behavioral outcomes. Such individual differences are likely to be important to consider in the design of optimal interventions.

The relationship identified between increased exposure to routines and increased physiological arousal following changes to those routines was also not clear cut, and was primarily driven by the differences in responses to changes to routines that had been established for 20 compared to 10 minutes. Taken together with the behavioral observation data, these data fit with a behavioral sequence model of temper outbursts. Previous research has demonstrated profiles of behaviors within a temper outburst, which progress in characteristic sequences (Oliver, Woodcock & Humphreys, 2009; Potegal, 2003; Green, Whitney & Potegal, 2011; Tunnicliffe, Woodcock, Bull, Penhallow & Oliver, 2014). Interestingly, behaviors more indicative of increased emotional arousal (such as emotional vocalizations or increased salivation) often occur together either preceding or following more challenging/disruptive behaviors. Thus, it is possible that the present increase in routine establishment from 10 to 20 minutes effected a small difference in the response to changes (relative to when routines had been established for longer) detectable primarily in the mean heart rate data; whereas with longer exposure to routines, larger responses to changes were observed in overt temper outburst behaviors, and the effect on mean heart rate was less pronounced. One important area for future research will be to better characterize the changes in physiological arousal that occur proceeding, during and following an outburst. This would be best achieved via conjunctive measurement from multiple indices of arousal such as heart rate, galvanic skin conductance and pulse rate.

However, the possible explanation for the observed relationship between routine establishment and emotional arousal discussed above must be considered tentatively. An important limitation with the present method for indexing emotional arousal must be noted. In order to control for the effects of movement on heart rate, activity data from a separate device that comprised an accelerometer were collected. These data suggested that participants' movement was not driving the main effect of exposure to routines on heart rate. However, they did not allow the effects of movement to be removed from the heart rate data. Portable devices are now available which contain an electrocardiogram alongside an accelerometer (e.g. Koehler, de Marees, Braun & Schaenzer, 2011), and these would provide a purer index of emotional arousal in future related studies.

Some of the primary limitations of the present study have already been highlighted. It is also pertinent to underline the potential limitation associated with the experimental setting used for the present study. Such a setting was chosen because it was reasoned; and has been demonstrated previously (e.g. Woodcock, Oliver & Humphreys, 2011); that environmental triggering events for challenging behaviors are less potent in experimental compared to natural settings. Thus, the study was designed to provide a stringent test of the present hypothesis. Using routines that were completely novel to participants before the study also allowed us to maintain the hypothesis test as stringent as possible. However, the need for studies examining longer and more realistic durations of routine establishment has already been highlighted. In addition, examination of these issues within a more natural environment may highlight other important factors, which were not evident in the experimental setting.

Finally, it is important to end with a note of caution about the generalizability of the present findings. There is fairly compelling evidence, discussed in the *Introduction*, that certain similarities exist across individuals with different neurodevelopmental disorders who show challenging behaviors following changes to routines or expectations (specifically with respect to the factors that immediately impact on those behaviors). However, the present study included participants with a single disorder. Future research is necessary to assess how far the findings of the present study are pertinent to individuals with other disorders. Notwithstanding this limitation, it is hoped that the present study will promote much needed systematic research that investigates the long term, prospective impact of environmental rigidity versus flexibility on individuals with neurodevelopmental disorders who show an elevated resistance to change.

References

Butler, J. V., Whittington, J. E., Holland, A. J., Boer, H., Clarke, D., & Webb, T. (2002). Prevalence of, and risk factors for, physical ill-health in people with Prader-Willi syndrome: a population-based study. *Developmental Medicine and Child Neurology, 44*(4), 248-255. doi: 10.1017/s001216220100202x

Buss, K. A., & Goldsmith, H. H. (1998). Fear and anger regulation in infancy: Effects on the temporal dynamics of affective expression. *Child Development, 69*(2), 359-374. doi: 10.1111/j.1467-8624.1998.tb06195.x

D'Cruz, A. M., Ragozzino, M. E., Mosconi, M. W., Shrestha, S., Cook, E. H., & Sweeney, J. A. (2013). Reduced Behavioral Flexibility in Autism Spectrum Disorders. *Neuropsychology, 27*(2), 152-160. doi: 10.1037/a0031721

Dimitropoulos, A., Feurer, I. D., Butler, M. G., & Thompson, T. (2001). Emergence of compulsive behavior and tantrums in children with Prader-Willi syndrome. *American Journal on Mental Retardation, 106*(1), 39-51. doi: 10.1352/0895-8017(2001)106<0039:eocbat>2.0.co;2

Donkin, L., Hickie, I. B., Christensen, H., Naismith, S. L., Neal, B., Cockayne, N. L., & Glozier, N. (2013). Rethinking the Dose-Response Relationship Between Usage and Outcome in an Online Intervention for Depression: Randomized Controlled Trial. *Journal of Medical Internet Research, 15*(10), 67-80. doi: 10.2196/jmir.2771

Ekas, N. V., Braungart-Rieker, J. M., Lickenbrock, D. M., Zentall, S. R., & Maxwell, S. M. (2011). Toddler Emotion Regulation With Mothers and Fathers: Temporal Associations Between Negative Affect and Behavioral Strategies. *Infancy, 16*(3), 266-294. doi: 10.1111/j.1532-7078.2010.00042.x

Ekman, P., Levenson, R. W., & Friesen, W. V. (1983). Autonomic nervous-system activity distinguishes among emotions. *Science, 221*(4616), 1208-1210. doi: 10.1126/science.6612338

Fernandez, C., Pascual, J. C., Soler, J., Elices, M., Portella, M. J., & Fernandez-Abascal, E. (2012). Physiological Responses Induced by Emotion-Eliciting Films. *Applied Psychophysiology and Biofeedback, 37*(2), 73-79. doi: 10.1007/s10484-012-9180-7

Fraser, J., Kirkby, K. C., Daniels, B., Gilroy, L., & Montgomery, I. M. (2001). Three versus six sessions of computer-aided vicarious exposure treatment for spider phobia. *Behavior Change, 18*(4), 213-223. doi: 10.1375/bech.18.4.213

Furniss, F., & Biswas, A. B. (2012). Recent research on aetiology, development and phenomenology of self-injurious behavior in people with intellectual disabilities: a systematic review and implications for treatment. *Journal of Intellectual Disability Research, 56*(5), 453-475. doi: 10.1111/j.1365-2788.2012.01534.x

Gomot, M., & Wicker, B. (2012). A challenging, unpredictable world for people with Autism Spectrum Disorder. *International Journal of Psychophysiology, 83*(2), 240-247. doi: 10.1016/j.ijpsycho.2011.09.017

Green, J. A., Whitney, P. G., & Potegal, M. (2011). Screaming, Yelling, Whining, and Crying: Categorical and Intensity Differences in Vocal Expressions of Anger and Sadness in Children's Tantrums. *Emotion, 11*(5), 1124-1133. doi: 10.1037/a0024173

Hollenstein, T., Granic, I., Stoolmiller, M., & Snyder, J. (2004). Rigidity in parent-child interactions and the development of externalizing and internalizing behavior in early childhood. *Journal of Abnormal Child Psychology, 32*(6), 595-607. doi: 10.1023/b:jacp.0000047209.37650.41

Holm, V. A., Cassidy, S. B., Butler, M. G., Hanchett, J. M., Greenswag, L. R., Whitman, B. Y., & Greenberg, F. (1993). Prader-willi syndrome - consensus diagnostic-criteria. *Pediatrics, 91*(2), 398-402.

Iellamo, F. (2001). Neural mechanisms of cardiovascular regulation during exercise. *Autonomic Neuroscience-Basic & Clinical*, *90*(1-2), 66-75. doi: 10.1016/s1566-0702(01)00269-7

Imeraj, L., Antrop, I., Sonuga-Barke, E., Deboutte, D., Deschepper, E., Bal, S., & Roeyers, H. (2013). The impact of instructional context on classroom on-task behavior: A matched comparison of children with ADHD and non-ADHD classmates. *Journal of School Psychology*, *51*(4), 487-498. doi: 10.1016/j.jsp.2013.05.004

Koehler, K., de Marees, M., Braun, H. & Schaenzer, W. (2011) Evaluation of two portable sensors for energy expenditure assessment during high-intensity running. *European Journal of Sport Science*, *13*(1), 31-41. doi: 10.1080/17461391.2011.586439

Kloosterman, P. H., Summerfeldt, L. J., Parker, J. D. A., & Holden, J. J. A. (2013). The obsessive-compulsive trait of Incompleteness in parents of children with autism spectrum disorders. *Journal of Obsessive-Compulsive and Related Disorders, 2*(2), 176-182. doi: 10.1016/j.jocrd.2012.11.004

Kuenssberg, R., Murray, A. L., Booth, T., & McKenzie, K. (2014). Structural validation of the abridged Autism Spectrum Quotient-Short Form in a clinical sample of people with autism spectrum disorders. *Autism, 18*(2), 69-75. doi: 10.1177/1362361312467708

Liu, J. H., Leung, P. W. L., McCauley, L., Ai, Y. X., & Pinto-Martin, J. (2013). Mother's environmental tobacco smoke exposure during pregnancy and externalizing behavior problems in children. *Neurotoxicology, 34*, 167-174. doi: 10.1016/j.neuro.2012.11.005

Lopez, B. R., Lincoln, A. J., Ozonoff, S., & Lai, Z. (2005). Examining the relationship between executive functions and restricted, repetitive symptoms of Autistic Disorder. *Journal of Autism and Developmental Disorders*, *35*(4), 445-460. doi: 10.1007/s10803-005-5035-x

Lunkenheimer, E. S., Albrecht, E. C., & Kemp, C. J. (2013). Dyadic Flexibility in Early Parent-Child Interactions: Relations with Maternal Depressive Symptoms and Child Negativity and Behavior Problems. *Infant and Child Development, 22*(3), 250-269. doi: 10.1002/icd.1783

Martin, N., Oliver, C. & Hall, S. (1998). *ObsWin: Software for the collection and analysis of observational data.* Birmingham: Univeristy of Birmingham.

Meddis, R. (1975). *Statistical Handbook for Non-statisticians*. McGraw-Hill Book Company: Maidenhead, Berkshire, England.

Mesibov, G. B., & Shea, V. (2010). The TEACCH Program in the Era of Evidence-Based Practice. *Journal of Autism and Developmental Disorders, 40*(5), 570-579. doi: 10.1007/s10803-009-0901-6

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100. doi: 10.1006/cogp.1999.0734

Monsell, S., & Mizon, G. A. (2006). Can the task-cuing paradigm measure an endogenous task-set reconfiguration process? *Journal of Experimental Psychology-Human Perception and Performance, 32*(3), 493-516. doi: 10.1037/0096-1523.32.3.493

Moss, J., Oliver, C., Arron, K., Burbidge, C., & Berg, K. (2009). The prevalence and phenomenology of repetitive behavior in genetic syndromes. *J Autism Dev Disord, 39*(4), 572-588.

Nessler, D., Friedman, D., & Johnson, R. (2012). A new account of the effect of probability on task switching: ERP evidence following the manipulation of switch probability, cue informativeness and predictability. *Biological Psychology, 91*(2), 245-262. doi: 10.1016/j.biopsycho.2012.07.005

Oliver, C., Woodcock, K. A., & Humphreys, G. W. (2009). The Relationship between Components of the Behavioral Phenotype in Prader-Willi Syndrome. *Journal of Applied Research in Intellectual Disabilities, 22*(4), 403-407. doi: 10.1111/j.1468-3148.2008.00475.x

Page, E. B. (1963). Ordered hypotheses for multiple treatments - A significance test for linear ranks. *Journal of the American Statistical Association, 58*, 216-230.

Ponitz, C. C., Rimm-Kaufman, S. E., Grimm, K. J., & Curby, T. W. (2009). Kindergarten Classroom Quality, Behavioral Engagement, and Reading Achievement. *School Psychology Review, 38*(1), 102-120.

Potegal, M., & Davidson, R. J. (2003). Temper tantrums in young children: 1. Behavioral composition. *Journal of Developmental and Behavioral Pediatrics, 24*(3), 140-147.

Rainville, P., Bechara, A., Naqvi, N., & Damasio, A. R. (2006). Basic emotions are associated with distinct patterns of cardiorespiratory activity. *International Journal of Psychophysiology, 61*(1), 5-18. doi: 10.1016/j.ijpsycho.2005.10.024

Richards, C., Oliver, C., & Allen, D. (2010). The function of self-injurious behavior in autism spectrum disorder. *Journal of Applied Research in Intellectual Disabilities, 23*(5), 431-431.

Russo, N., Flanagan, T., Iarocci, G., Berringer, D., Zelazo, P. D., & Burack, J. A. (2007). Deconstructing executive deficits among persons with autism: Implications for cognitive neuroscience. *Brain and Cognition, 65*(1), 77-86. doi: 10.1016/j.bandc.2006.04.007

Sabaratnam, M., Murthy, N. V., Wijeratne, A., Buckingham, A., & Payne, S. (2003). Autistic-like behavior profile and psychiatric morbidity in Fragile X Syndrome - A prospective ten-year follow-up study.*European Child & Adolescent Psychiatry, 12*(4), 172-177. doi: 10.1007/s00787-003-0333-3

Sinha, N., Manohar, S., & Husain, M. (2013). Impulsivity and apathy in Parkinson's disease. *Journal of Neuropsychology, 7*(2), 255-283. doi: 10.1111/jnp.12013

Sparrow, S.S., Chicchetti, D.V. & Balla, D. (2005). *Vineland Adaptive Behavior Scales (Second Edition).* Minneapolis, MN: Pearson Assessment.

Tanimura, Y., Yang, M. C., & Lewis, M. H. (2008). Procedural learning and cognitive flexibility in a mouse model of restricted, repetitive behavior. *Behavioral Brain Research, 189*(2), 250-256. doi: 10.1016/j.bbr.2008.01.001

Tunnicliffe, P., Woodcock, K., Bull, L., Oliver, C., & Penhallow, J. (2014). Temper outbursts in Prader-Willi syndrome: causes, behavioral and emotional sequence and responses by carers. *Journal of Intellectual Disability Research, 58*(2), 134-150. doi: 10.1111/jir.12010

Walz, N. C., & Benson, B. A. (2002). Behavioral phenotypes in children with Down syndrome, Prader-Willi syndrome, or Angelman syndrome. *Journal of Developmental and Physical Disabilities, 14*(4), 307-321. doi: 10.1023/a:1020326701399

White, S. J. (2013). The Triple I Hypothesis: Taking Another('s) Perspective on Executive Dysfunction in Autism. *Journal of Autism and Developmental Disorders, 43*(1), 114-121. doi: 10.1007/s10803-012-1550-8

Whittington, J., Holland, A., Webb, T., Butler, J., Clarke, D., & Boer, H. (2004). Cognitive abilities and genotype in a population-based sample of people with Prader-Willi syndrome. *Journal of Intellectual Disability Research, 48,* 172-187. doi: 10.1111/j.1365-2788.2004.00556.x

Woodcock, K. A (2008). Mapping a pathway from genes to behavior in Prader-Willi syndrome. *PhD thesis,* University of Birmingham, UK.

Woodcock, K., Oliver, C., & Humphreys, G. (2009). Associations between repetitive questioning, resistance to change, temper outbursts and anxiety in Prader-Willi and Fragile-X syndromes. *Journal of Intellectual Disability Research, 53*, 265-278. doi: 10.1111/j.1365-2788.2008.01122.x

Woodcock, K. A., Humphreys, G. W., Oliver, C., & Hansen, P. C. (2010). Neural correlates of task switching in paternal 15q11-q13 deletion Prader-Willi syndrome. *Brain Research, 1363*, 128-142. doi: 10.1016/j.brainres.2010.09.093

Woodcock, K. A., Oliver, C., & Humphreys, G. W. (2009). Task-switching deficits and repetitive behavior in genetic neurodevelopmental disorders: Data from children with Prader-Willi syndrome chromosome 15 q11-q13 deletion and boys with Fragile X syndrome. *Cognitive Neuropsychology, 26*(2), 172-194. doi: 10.1080/02643290802685921

Woodcock, K. A., Oliver, C., & Humphreys, G. W. (2011). The relationship between specific cognitive impairment and behavior in Prader-Willi syndrome. *Journal of Intellectual Disability Research, 55*, 152-171. doi: 10.1111/j.1365-2788.2010.01368.x

Supplementary methods

Semi-structured interview schedule (Telephone)

Temper outburst/tantrum –

Highly emotional response. Period of crying, screaming, angry ranting, shouting, stamping feet, or kicking. Can last for a prolonged period of time.

1. Does ___________ ever display temper outburst behaviour?

2. If so, what behaviours does __________ show during a typical episode?

3. How often do the temper outbursts occur?

4. Think about the last time a temper outburst occurred, what seemed to trigger the behaviour in this example?

5. In the example you thought about, how did you respond to the temper outburst?

6. Roughly, how many times does the trigger you mentioned actually result in a temper outburst?

7. Are there times when this particular trigger does not actually trigger a temper outburst?

8. Roughly, how many times after a temper outburst would you respond in the way that you mentioned in the example?

9. What behaviours does __________ typically show after an outburst?

10. Think of other examples in which a temper outburst occurred, what seemed to trigger this and how did you respond?

11. How long roughly do the temper outbursts last for?

Routines –

Actions or procedures that are followed regularly, often repetitiously.

12. Does _________ have particular routines?

13. Are these routines important to ________?

14. What happens if the routine changes?

Review

15. Finally go over behaviours, antecedents and consequences listed.
    Anymore behaviours, antecedents, consequences?

Can these antecedents and consequences be categorised? Do this with parent/carer.

Description of the change challenge games

The games

- Mountain dice
- Colour and shape
- Drawing dice game
- Counter game

General points on the games and changes

- There should first be a familiarisation period of time where the participant is taught all of the games and understands the rules. This should last about 20mins, 5mins spent on each game
- The game is then played for the pre-set amount of time.
- When this time is up there is a 5 minute period where changes are introduced.
- There are a number of changes available. These should be introduced in a random order/as the experimenter chooses. It is possible to pick the changes in order to help the participant!
- In each 5min change there should be 5 changes in total and no change should be repeated.
- After each change "go back to the old rules".
- Changes should be introduced clearly e.g. "now we're going to pick up two cards".
- If changes are questioned can say "that's how we're playing it now".
- Changes should be kept for 1go. Then the game should go back to the previous rules.
- Helps if continue to talk through the process so that the participant does not get confused.
- Try not to give a reason for the change- there is no need to justify them
- After the final change play for a while with the "old" rules before the 5 minutes are up
- At the end of each game ask the player to fill in the self report sheet

Mountain dice game

Equipment

- 1 die
- Laminated mountain sheets and people to move
- Whiteboard markers and cloth/tissue

Rules

- The object of the game is to climb and descend the 3 mountains, in number order. Moving from left to right (i.e. starting from the smallest mountain onwards to the largest- *demonstrate by moving person*)
- Each player gets a strip with the three mountains on.
- Players take it in turns to roll a dice. To move up and down the mountains the players need to cross the numbers off in order.
- They can cross a number off when they roll that number on the dice.
- The numbers need to be crossed off in order so the player needs to roll a 1 to cross off their first number and move their man up.

Changes

- Climb the mountain backwards. I.e. move their person to the bottom right of the mountain "Now for this go we're going to move the person here and try to roll a ** to try to climb the mountains backwards".
- Introduce an extra die so we are now using two dice. "So for this go we will roll two dice and you need to try to get a *". Talk them through their roll. E.g. if needed 5 and rolled 5 1 can cross 5. If rolled 5 & 6 can cross both off.
- Cross the numbers off in any order "This turn we're going to cross off whatever number you roll, it doesn't have to go in order, so have a go. Ok, now cross off all your *s"
- Roll again – "This time we're going to have 2 turns"
- Start at top of a new mountain- Move their person to top of largest mountain "For this go we're going to put your person here and try to roll a 6"
- Old rules – cross off in order, roll only one die, go left to right.

Colour and shape card game

<u>Equipment</u>

- "my first colour and shape snap cards"
- Crayons
- Shape worksheets

<u>Rules</u>

- Players need to colour in the shapes in order (1 first, then 2 etc)
- Players take it in turns to pick a card from the pack
- If they pick the right shape they can colour it in the same colour as on the card
- So when they pick a rectangle they can colour this in with the same colour as on the card.

<u>Changes</u>

- <u>Ignore the order</u>, "This time we are going to ignore the order in which we colour the shapes in, you don't have wait to draw the * this time you can colour the shapes in any order. Pick a card. Okay now you can colour that!"
- <u>Follow the colour</u> "This time we're ignoring the shape that is on the card but following the colour. So look at your card – what's the colour? Colour in your next shape in that colour"
- <u>Follow the shape</u> (jump to that shape) "This time we're using a different colour than on the card. So choose any colour that's not ****"
- <u>Take extra goes</u> "This time we get two rolls"
- <u>Jump to the end</u> "So now we have to try to colour the oval"
- <u>Old rules:</u> in order, follow shape and colour, only one card drawn.

Drawing Dice Game

Equipment

- Dice
- Worksheets
- Pens/crayons

Rules

- Each number is assigned a body part
- Players take it in turns to roll the dice until they get the right number to draw the next body part
- Players have to get the numbers in order to draw
- So when roll a 1 they can draw the head etc
- Cross off each number when it has been drawn (to prevent confusion)
- When animal is finished they can pick another animal

Changes

- Draw feet first "This roll we're going to try to roll a 6 for feet"
- Any order "This time we don't have to roll the dice in order. So roll. Now you can draw ***"
- Allow extra rolls. "This time we're going to roll twice each"
- Introduce 2 dice. "This time we're going to roll two dice" (talk them through)
- Roll the dice for each other "This time we are going to roll the dice for each other" (talk them through)
- Old rules, draw body parts in order

Counter Game

Equipment

- Counters
- Pot
- Cards
- Dice

Rules

- Lay out the cards in a heap on the table
- Divide the counters between the players evenly
- Take it in turns to roll the dice
- Find a card with the same number of dots
- Turn the card over and put that number of counters into the pot
- The card must be put in a "used" pile
- If players cannot find a card with that number on (because all cards with that number on are in the "used" pile) then their turns end without putting any counters in the pot

Changes

- Roll again "This time we're going to have two rolls each"
- Put cards back (not in used pile) "This time we're not going to put the cards in the used pile. We'll shuffle them back in"
- Match colour AND number on back "for this go we're going to look at the number and COLOUR on the back. OK, so you've got a **** so put in (number) (colour) counters".
- Follow the colour- ALL counters "This go we're just going to look at the COLOUR on the back and put in all our counters of that colour. So you've got *** so put in all your *** counters".
- Don't turn over "This time we're not going to turn over the card but just look at the front and put in the same number of counters as dots".
- Old rules : Reinstate the "used pile", follow only number on back and ignore colours

Supplementary analyses and results

Individual participant level behaviour observation data

Since the present design was repeated measures and substantial variability in observed behaviour was demonstrated across participants, the percentage of temper outburst behaviours shown during Disrupt conditions following Establish conditions of 10, 20, 40 and 80 minutes was examined at an individual participant level (*Figures S1, S2 and S3*). Overall 10 of the 16 participants demonstrated a pattern of behaviour broadly consistent with the group profile presented in the main manuscript.

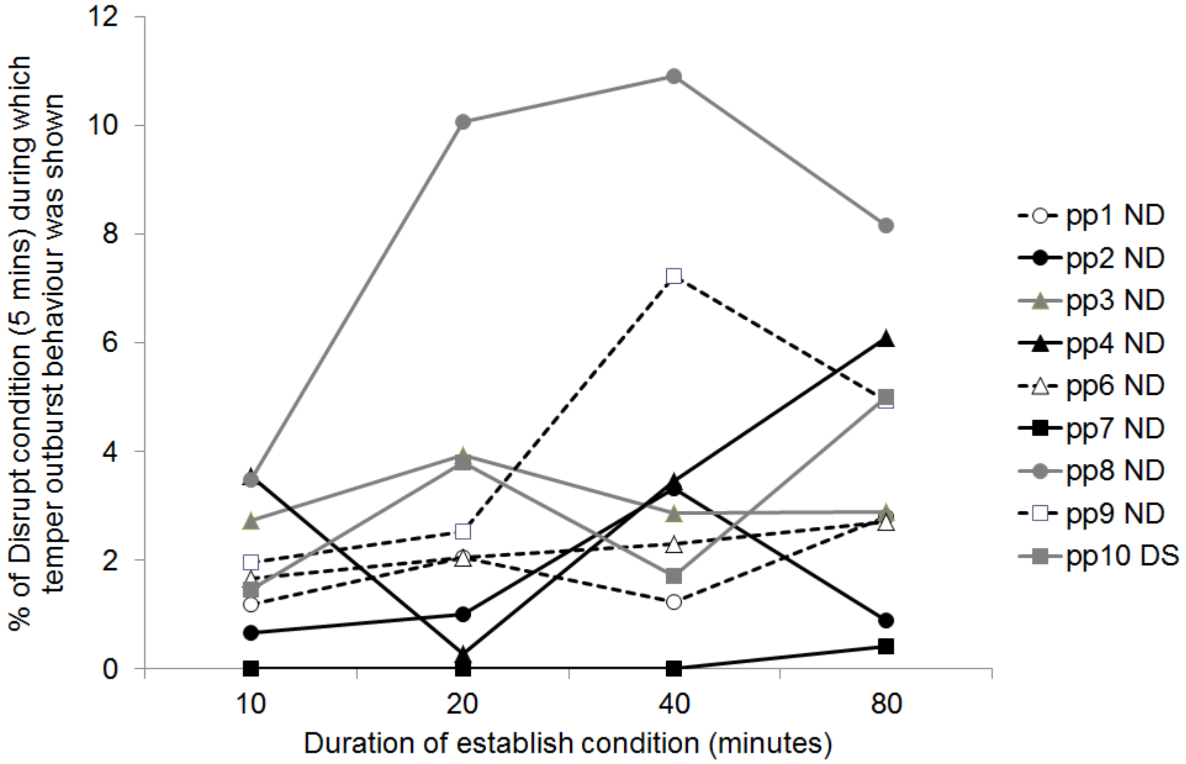

Figure S1 The percentage of each Disrupt condition (5 minutes, following different lengths of Establish conditions) during which each participant demonstrated temper outburst behaviours (participants are numbered arbitrarily for the purpose of display). The figure *groups 9 of the participants with a pattern of behaviour broadly consistent with the group profile.*

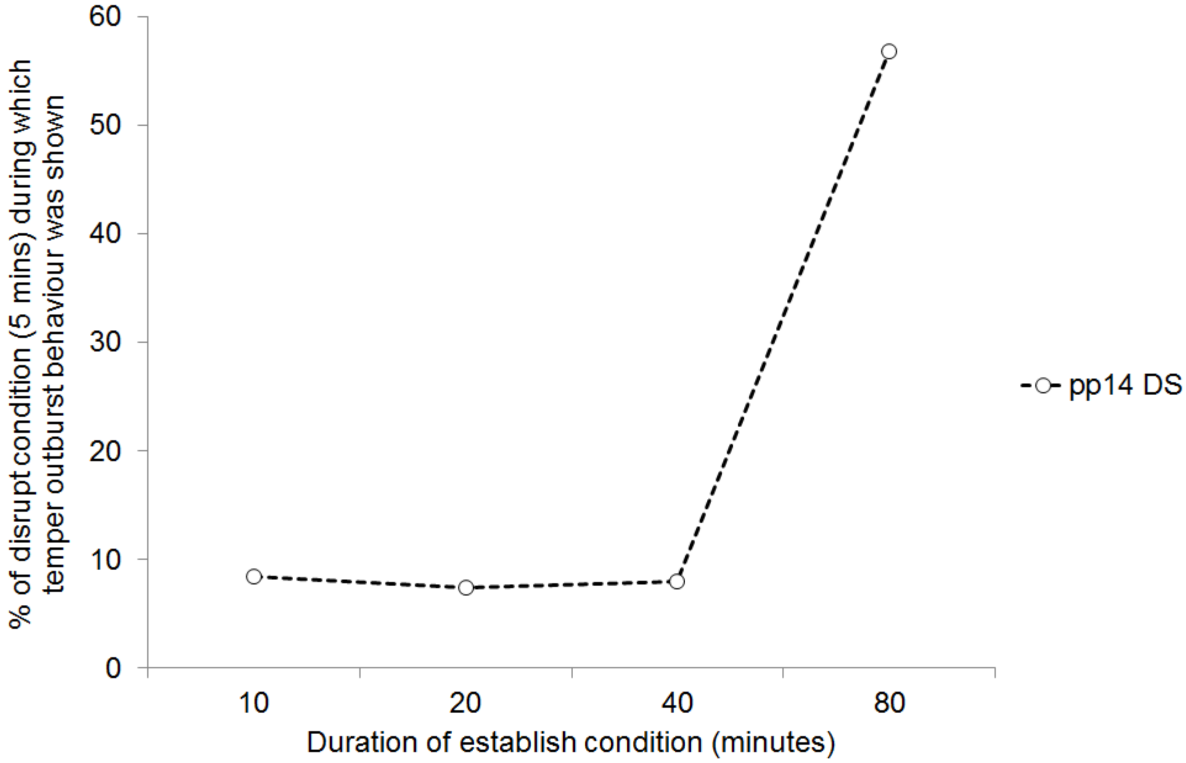

Figure S2 The percentage of each Disrupt condition (5 minutes, following different lengths of Establish conditions) during which one participant demonstrated temper outburst behaviours (participants are numbered arbitrarily for the purpose of display). The figure shows the tenth participant with a pattern of behaviour broadly consistent with the group profile.

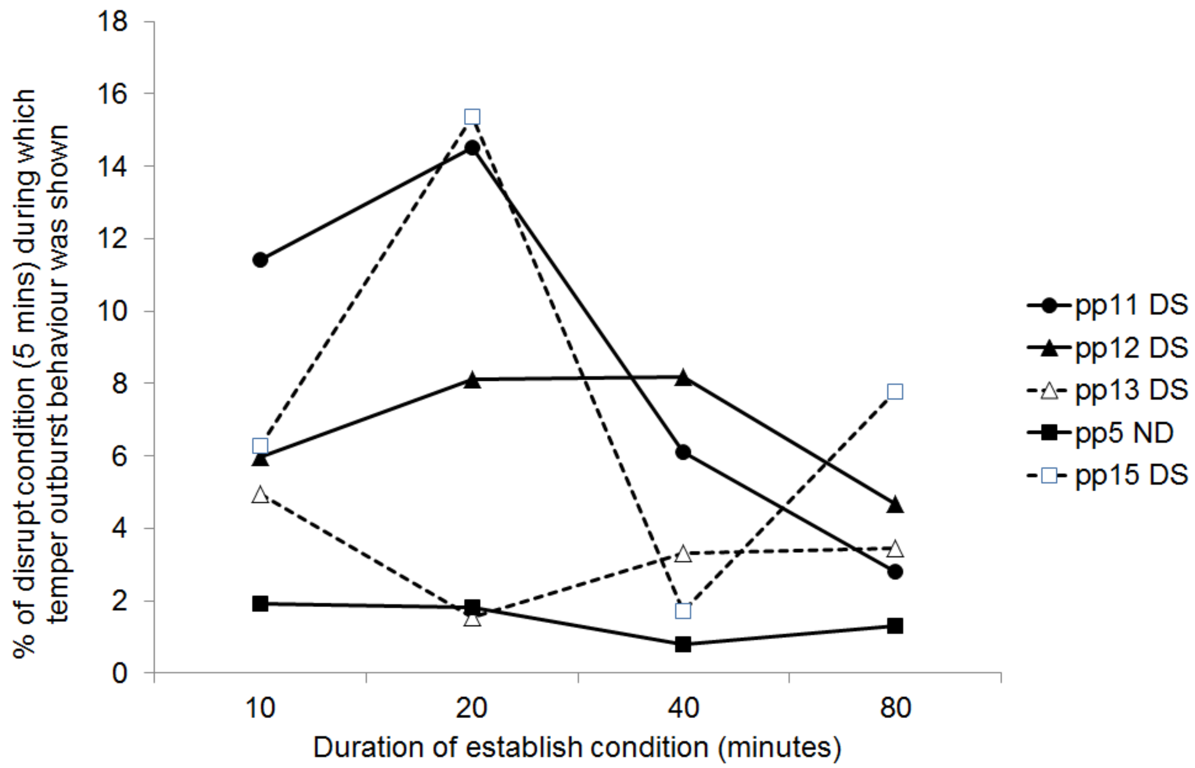

Figure S3 The percentage of each Disrupt condition (5 minutes, following different lengths of Establish conditions) during which one participant demonstrated temper outburst behaviours (participants are numbered arbitrarily for the purpose of display). The figure shows the six participants with a pattern of behaviour that was not broadly consistent with the group profile.

Exploratory analyses of behaviour observation data

Following initial analysis, further coding was introduced. This was based on the observation that whilst playing the games, some participants appeared to be distracted at times. It was reasoned that given that the design of the study relied on individuals being aware when changes had occurred, any time periods when individuals were not paying attention to the games may affect results.

Distraction was thus operationally defined as any period of time when the participant was not engaged or paying attention to the game. This included times when: the participant did not notice when it was their turn; did not look at the game materials as they took their own turn, and did not look at the researcher or the game when the researcher took their turn; talked about something unrelated to the game (unless the participant was watching and paying attention to the game whilst talking); carried out an action unrelated to the game (e.g. looks in a book), which could include actions with game items (e.g. smelling the counters); looked/talked to someone else in the room or looked at the camera/out the window, but this excluded brief (up to 3 seconds) glances away from the researcher/the game; played with the physiological recording equipment; asked questions or made comments signifying boredom (e.g. "do we have long left?" etc).

Inter-rater reliability was assessed on 25% of each participant's distraction coding and good inter-rater reliability was established, with a kappa value of .66.

Exploratory results on behaviour observation data

Results are reported in the main manuscript relating to participants who demonstrated relatively less distraction (did not demonstrate distraction for more than 20% of the duration of any of the Disrupt conditions; n=9; *Figures S4 and S5*). Distraction being shown for 20% of the duration of at least one Disrupt condition was treated as an arbitrary cut-off in this exploratory analysis and was based on the traditional confidence limits reported for statistical tests in Psychology research. In Figures S1-S3, participants falling below the 20% cut-off for distraction are labelled with NS (not distracted) and those falling above the 20% cut-off are labelled with DS (distracted). As demonstrated from these figures, only two participants with responding broadly consistent with the group profile fell into the *distracted* group, whereas four out of five participants with responding inconsistent with the group profile fell into the *distracted* group.

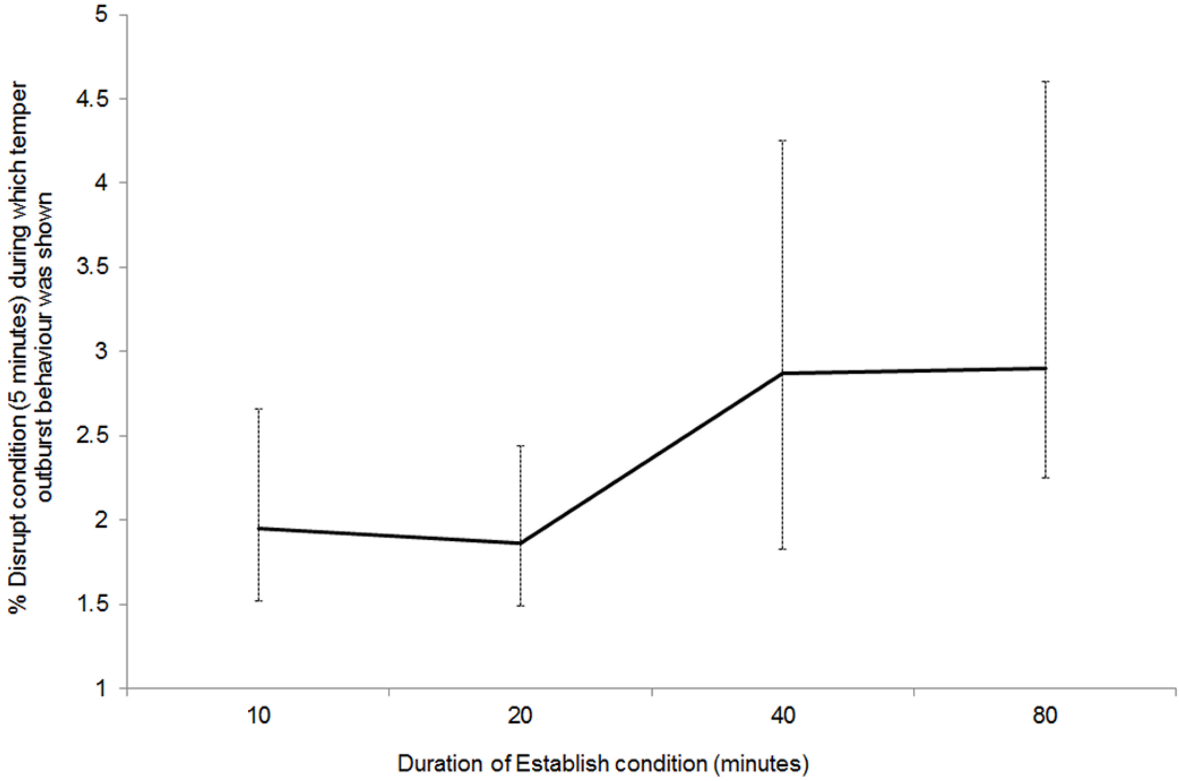

Figure S4 The median percentage of five minute Disrupt conditions during which temper outburst behaviours were demonstrated by participants (n=9) who displayed distraction during less than 20% of all Disrupt conditions. Horizontal dashed lines represent the inter-quartile range.

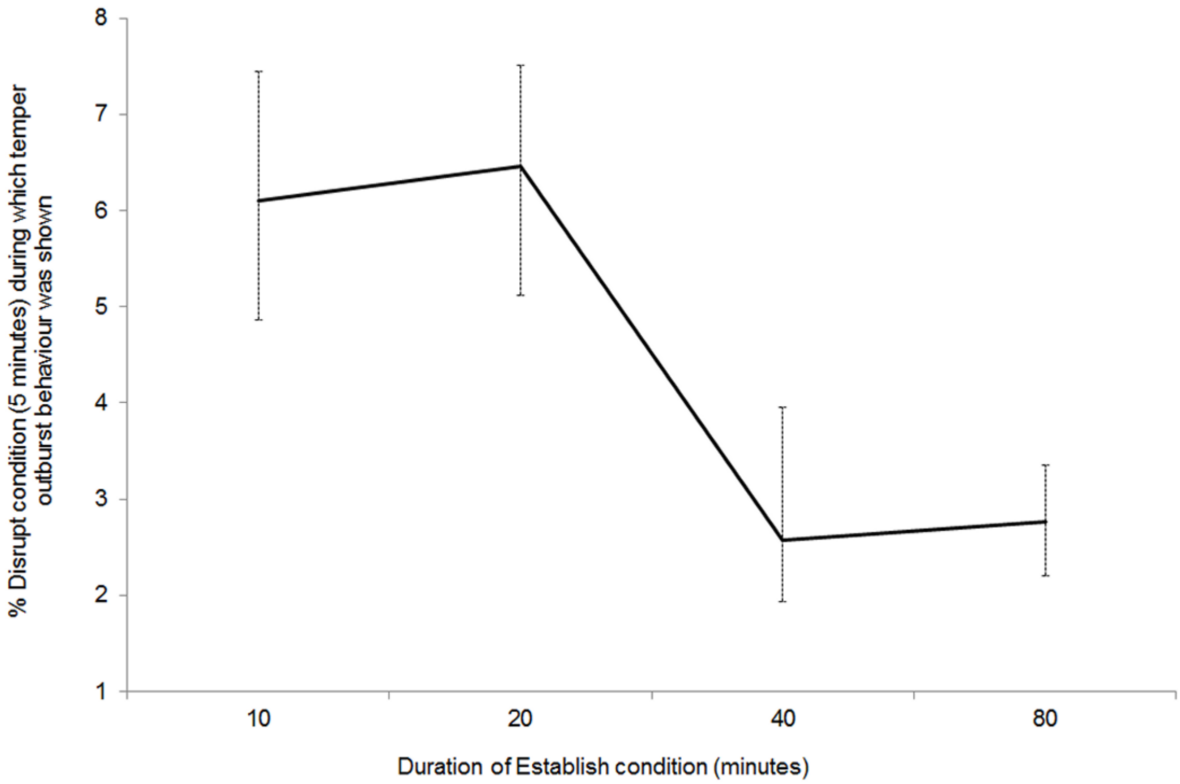

Figure S5 The median percentage of five minute Disrupt conditions during which temper outburst behaviours were demonstrated by participants (n=5) who displayed distraction during at least 20% of at least one Disrupt condition. Horizontal dashed lines represent the inter-quartile range.

The mean percentage of all Disrupt conditions during which the participants in the relatively less distracted group demonstrated temper outbursts was 2.9%; and the corresponding mean for the relatively more distracted group was 5.2%. This difference between the groups bordered significance (*Wilcoxon Man-U:* 43, *p*= .066). A Spearman's correlation analysis demonstrated a significant positive association between the percentage of Disrupt conditions during which temper outburst behaviours were shown and the percentage of time spent participants were distracted in the corresponding conditions (r = 0.52, p = .05; *Figure S6*). There was an overlap in the definitions of distraction and temper outburst behaviour because ignoring requests could potentially be comprised in either category. Thus, incidents of ignoring requests were excluded from this analysis

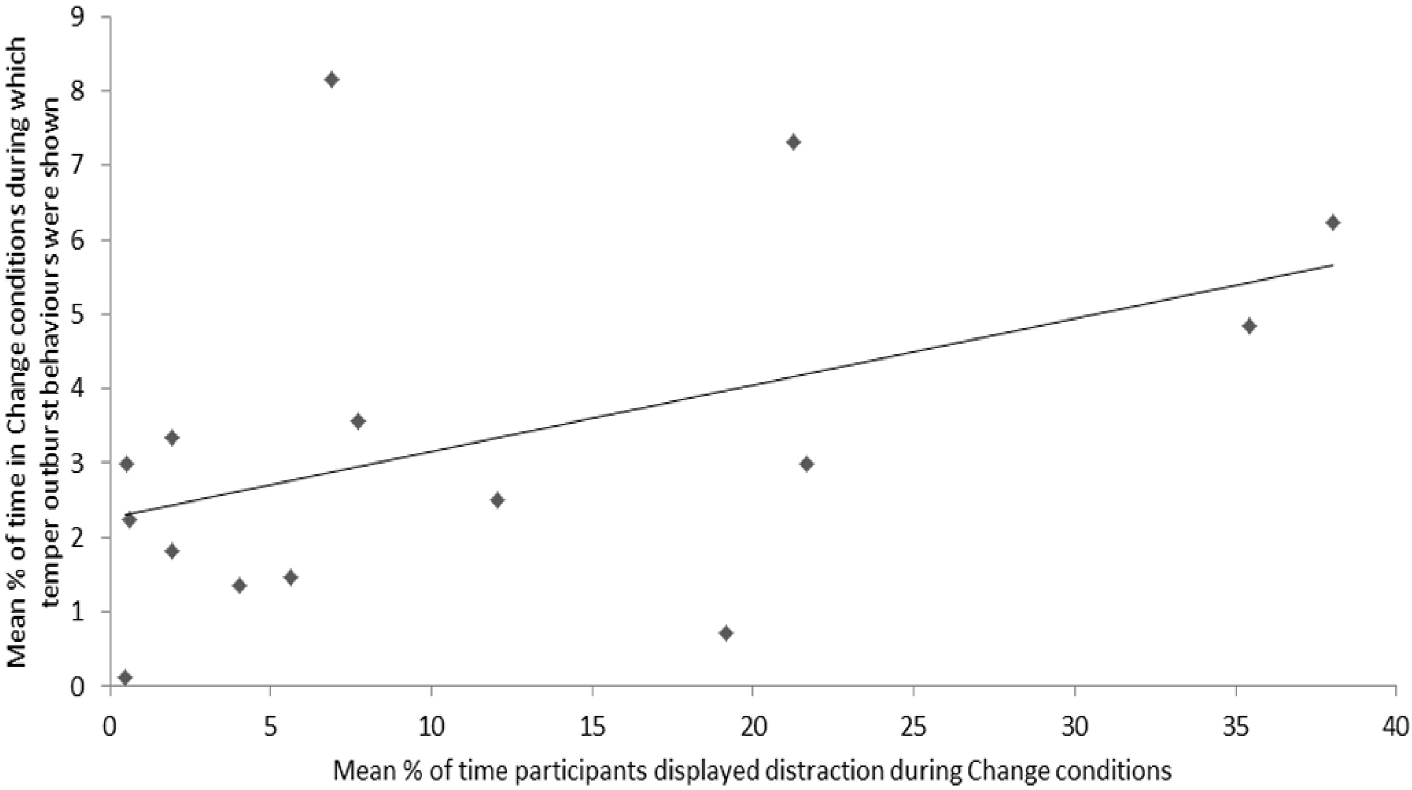

Figure S6 Relationship between the mean percentage of time participants were distracted and the mean percentage of time particpants displayed temper outburst related behaviours in Disrupt conditions.